

Authors Response to Reviewer #1

We appreciate your comments, which provided important means for improving the manuscript. In the following we address them one by one.

Reviewer#1: What's not as clear, or not taken as further here, is the relevance for furthering our understanding / reducing uncertainty in the impacts of aerosols on health and climate

Authors: We agree. We added the following to the end of the conclusion: "As we show that assimilation can improve estimates of surface PM_{2.5}, this technique can be used to generate analysis with high temporal and spatial resolution for use in health assessments (e.g., Silva et al., 2013). Also, the improved aerosol loads can help to better estimate aerosol climate forcing. Finally, the assimilation can be used in forecasting mode to predict air quality more accurately."

Reviewer#1: My broadest request would be to include additional discussion of the implications of the work here in terms of reducing uncertainty in knowledge of aerosol sources, and how 3D-Var, while not explicitly designed to adjust emissions, can be interpreted in that light. The authors do in fact touch lightly on several such issues (issues with NO_x or SO₂ inventories, or dust and sea salt concentrations in the boundary conditions being overestimated); such aspects could be brought out a bit more.

- A problem with the approach is that the fundamental source of model error, namely emissions, is not improved by the assimilation. This problem is most visible in plots such as Fig 6, where the observations only briefly pull down the model to values in better agreement with observations, only to pop right back up again as soon as the impact of the assimilation has subsided. Fundamentally, this problem required a different assimilation approach, or if keeping with 3D-Var perhaps the error correlation length / time scales need reconsideration.

Authors: The technique can be used to modify emissions. This can be done by further analysis of the increments. However, in our study most of the PM_{2.5} and AOD overestimation comes from overestimation of dust (check Fig. 4, the scale of "other" and PM_{2.5} is way higher than for the rest of the species) which is tracked back to the dust boundary conditions being too high. Thus, adjusting just emissions would end up reducing them when they do not necessarily need to be reduced (e.g., case of SO₄). Thus, future approaches would have to improve emissions and boundary conditions simultaneously in order to account for all uncertainties. The following was added:

(after 2nd paragraph section 3.1) As issues with local emissions are found, an emission inversion along with data assimilation could be performed as suggested by other studies (Jiang et al., 2013). However, as the major problem in this study arises from the dust boundary conditions, adjusting just emissions would end up reducing them when they do not necessarily need to be reduced (e.g., case of SO₄). Thus, future studies performing data assimilation and emissions inversions, would also need to assimilate chemical boundary conditions.

(at the end of the 1st paragraph of the conclusion) This is all demonstrated on a 3DVAR assimilation system, but it could eventually be applied in more sophisticated frameworks such as 4DVAR or Kalman filter systems to make use their strengths over 3DVAR (e.g., Pagowski and Grell, 2012). These methods would allow performing data assimilation simultaneously with boundary conditions and emissions inversions (e.g., Elbern et al., 2007), which is likely to extend aerosol improvements further on time.

- Abstract: Possible to add some quantitative aspects of the results to the abstract? At the moment the description is all qualitative.

Authors: Some quantitative results were added to the abstract.

- Intro: The rationale for using 3D-Var to address model uncertainty needs to be made more clear, or the introduction could use revision to focus more directly on the question of aerosol forecasting and the importance of this work in that light.

Authors: The use of the 3DVAR method over others is not the main topic of the paper, so we think this discussion does not belong in the introduction. On this point, the following is already stated in the conclusions of the ACPD manuscript: "This is all demonstrated on a 3DVAR assimilation system, but it could eventually be applied in more sophisticated frameworks such as 4DVAR or Kalman filter systems to make use their strengths over 3DVAR (e.g. Pagowski and Grell, 2012)." The following was added to the "Assimilation system" subsection (2.2) first paragraph to explain the rationale for using 3D-Var: "Even though more sophisticated assimilation schemes such as 4DVAR (Benedetti et al., 2009) and Ensemble Kalman filter (Pagowski and Grell, 2012) can be used for assimilation, we chose 3DVAR as a computationally inexpensive but powerful way to demonstrate AOD assimilation for the MOSAIC aerosol scheme, without having to perform an ensemble of simulations or develop the WRF-Chem adjoint."

- 12216.10: Also, Wang et al. (GRL, 2012) and Xu et al. (JGR, 2013), constrain emission using 4D-Var assimilation of AOD.

Authors: Thanks for the references, these were added.

- 12220: Is choice of the form of the control parameter or observation (linear or log scale) more or less consistent with the implicit assumption in using Eq (1) that x and y are normally distributed?

Authors: Good point. By using log parameters we are assuming that the errors on the log of the parameters are normally distributed (or that the errors are log-normal). As aerosol concentrations and AOD are positive, then it is likely that the errors are of multiplicative nature. Added the following to that

paragraph: “As both aerosol concentration and AOD are positive, it is likely that their errors are of multiplicative nature, and the use of a transformation becomes more natural as Eqn. 1 implicitly assumes that the errors are normally distributed (Bocquet et al., 2010).”

- 12220: I follow the explanation of how assimilating concentration vs mass will be different, but the reason for preferring the latter formulation hasn't been explained. Only the consequences of using concentrations are mentioned, and it isn't obvious to me why these consequences would be undesirable. To improve estimates of aerosol on climate, wouldn't we want to target grid cells that have the largest impact on the column AOD?

Authors: On that explanation, one assumption is that the uncertainty is the same for both cells. If the correction is not made, two cells with equal uncertainty and concentration will not be changing in the same way. As explained in the text, our resolution is very fine close to the ground (~50m) compared to the uppercells (up to 800m), so when this correction was not applied, changes in stratospheric aerosol that were not realistic were made, and surface concentrations were barely modified. Concentrating the changes in gridcells with the largest impact on AOD would be biased, as we know there could be errors in all heights, modifications it will end up depending on the model configuration rather than in the uncertainty.

A sentence was added to the text to better explain the consequences: “For two given grid-cells in the same column and containing the same aerosol concentrations and uncertainty, the grid with the deeper thickness will contain higher sensitivities, as the same change in concentration will generate a higher increase in AOD due to the deeper layer. **This will end up in the assimilation preferentially modifying concentrations in those deeper gridcells, biasing the model.** By multiplying by the thickness, we avoid the assimilation favoring changes in deeper grid-cells, which could be important in configurations with great vertical variability as the one used in this study.”

- 12221.06: Perhaps it will be discussed further later, but it's not clear from here how these values are chosen, or what their uncertainty is, or what their impact on the results might be.

Authors: Text was added to better explain what they mean, how they are chosen and what impact they have. These parameters are just for handling upper and lower bounds and they are chosen by the user, so we don't think that an uncertainty analysis of them is necessary. The text was modified in the following way: “... and use the parameters k_{uc} and k_{lc} for weighting the constraint, with higher values giving a higher weight to the terms in Eq. (1), thus allowing a smaller departure of x from the target bound once it has been exceeded. x_{uc} and x_{lc} represent the desired bounds for the control variable and are calculated as multiplicative factors applied to the prior (additive in the case of LN control variable). In the experiments, x_{uc} and x_{lc} were chosen equal to $5 \cdot x_b$ and $0.01 \cdot x_b$, meaning that the upper and lower bound terms are activated during minimization when x is over 5 times or below 1/100 times the background, respectively. The weights of the constraint term k_{uc} and k_{lc} were equal to 0.5 and 0.05,

which were chosen experimentally by trying different values and keeping a range that both restricts x to the bounds and at the same time keeping the constraint term from becoming the largest term in the functional J . Higher weight and more constrained multiplicative bound are given for the upper constraint as we found that overly increasing concentrations (i.e. incorrectly high AOD retrieval) can excessively damage the forecasts.”

- 12223.14: Could it be explained what are “tangent linear and adjoint tests”?

Authors: The following was added. “The tangent linear (TL) and adjoint of this code were obtained using the automatic differentiation tool TAPENADE v 3.6 (Hascoët and Pascual, 2004). Two tests were performed to validate the code generated. First, the TL code was tested using the TL test, which consists of comparing the derivatives obtained from the code against finite differences using the forward code, obtaining better agreement as the perturbation used was reduced, which is considered a successful test. Second, the adjoint code was tested using the adjoint test, which consists in generating derivatives with the TL code and then using them as an input for the adjoint code. In this case, a successful test is obtained when, for different perturbations, the dot product of the derivatives generated with the TL is equal to machine precision to the dot product of the adjoint derivatives and the original perturbation (Zou et al, 1997), which was also accomplished.”

- 12222.19: Regarding the constant correlation of two size bins, does that mean that aerosol properties within the fine mode are equivalently correlated to two bins spanning the boundary of the coarse and fine modes? Is this physically reasonable?

Authors: Under the current implementation, yes, any two neighbor size bins will have an equivalent correlation with one another. Even if different correlation scales were going to be applied on different size bins, this will still not fix this problem, as bins located at the boundary of the fine and coarse bins would require being correlated anisotropically (e.g. the bin in the upper boundary of the fine mode, would have to be closely correlated with the smaller bin and show very low correlation with the upper bin). In the current implementation of recursive filters this would not be possible, as filters mimic a Gaussian correlation, thus they apply isotropic correlations (see Purser et al., 2003 for details). An implementation of this kind of correlation would require a huge amount of additional research which is out of the scope of this paper. We acknowledge this limitation with the following changes added to the first paragraph on Section “Background error covariance matrix”: “... By using recursive filters we incorporate the capacity to add correlations between aerosol size bins in GSI. Filter passes run along size bins in incremental order and are applied locally for each aerosol size distribution, in a similar way as vertical scales are applied (Wu et al., 2002). For simplicity, the inter size bin correlation length are specified in the namelist by the user and not computed through the method described in the next paragraph. However, we do not discard this possibility for future studies. The size bin correlation length scale was chosen equal to 2 bin units, which prevents excessive accumulation of innovations on a single size bin and distributes the changes along them. The isotropic nature of one-dimensional recursive

filters restricts the ability to apply different correlation scales to bins that have smaller and larger sizes than the reference one. Such anisotropic correlation would be preferred for bins located at the edges of fine and coarse distributions. We hypothesize this limitation could be partially overcome when computing the correlation with methods such as the one described next.”

- 12226.22: Regarding “increase as AOD is lowered” – wouldn’t the constant a term prevent this for small AOD?

Authors: No. Actually, the constant term is responsible for increasing the relative error as AOD is lowered. For instance, for $a=0.05$ and $b=0.15$ (over land error from Remer et al., 2005), if $\tau=1$, relative error is 20%, while if $\tau=0.1$ relative error is 65%, and if $\tau=0.01$ error is ~500%. This is one of the reasons why this error approach was not used.

- 12230.15: Are there any other works evaluating the dust simulations used here?

Authors: MOZART does not compute DUST online, and uses monthly means coming from the CAM model. Mahowald et al., (2006) evaluates dust simulations by comparing yearly AOD and deposition, so this validation might be too coarse for our purposes. The following was added to section 2.1: “MOZART uses monthly dust distributions from Community Atmosphere Model (CAM) (Mahowald et al., 2006) calculations, which are also used in this study.”

- 12232.17: It might be useful to reiterate here that you are discussing (I think) fractional error compared to surface PM_{2.5} at the AQS sites.

Authors: Added

- Mention Xu in intro?

Authors: It was added as emission inversion reference. It is not mentioned further as it is not an example of multi-wavelength or fine and coarse assimilation (Xu et al compute AOD at 0.65 μm which is further used in the inversion)

- 12234: It’s a bit hard to reconcile the discussion of the persistence of influence of the observations here with the results shown in Fig 6, where it appears that in locations like Trinidad Head and UCSB that the assimilation run relaxes back to the non-assimilation run often within a day or less.

Authors: We acknowledge the persistence of the influence of observations, but we clearly say that error reductions are smaller for the 21 hour forecast. For the case of Trinidad Head, it is a coastal site on northern California (see Fig 1) and follows the same description for “over ocean 125W to the west”, where we clearly say that the assimilated model goes back to the non-assimilated one. Also, further down on the same paragraph this is mentioned for the Trinidad head station. In the case of USCB, with the exception of a couple of days, the assimilated does not go completely back to the non-assimilated, so small but positive error reductions are still found (which is described in the text). The case of consistency with La Jolla station is also mentioned. Some modifications were made to that paragraph.

- 12215.23: First sentence is a bit awkward and could use a bit of work. I think it’s the “play multiple roles including” part that is odd. Suggest something like “aerosols interact with society and the environment in several important ways –”

Authors: Thanks, this was replaced.

- 12217.6: Not sure what is meant by “performed over models”

Authors: Changed to: “However, assimilation performed for aerosol treatments that have higher degrees of freedom (i.e., multiples species and multiples size bins) may be useful when assimilating many data sources at the same time, as both the total mass and aerosol size distribution could be modified to produce a better fit to observations.”

- 12224.4: bins “of” MOSIAC?

Authors: MOSAIC has two options in WRF-Chem, the 4 and 8 bins treatments. In this study, we used the 8 bins, but everything was coded for both. To avoid misunderstandings, we erased the 4 bin part.

Changed to: “For simplicity, we consider fine mode as aerosols with a dry diameter equal or less than 625 nm (first 4 size bins from the 8 bins of MOSAIC), which is in agreement with the cut-off diameter of ...”

All the other editorial comments were included as requested.

Authors Response to Reviewer #2

We appreciate your comments, which helped improve the manuscript and broaden its impact. In the following we address them one by one.

Reviewer#2: I would recommend publication in its present form with perhaps a slight change in title from “Aerosol optical depth assimilation for a size-resolved sectional model: impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale forecasts” to “Aerosol optical depth assimilation for a size-resolved sectional model: impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale analyses and forecasts”

Authors: We agree, the title has been changed to what the reviewer suggests.

Reviewer#2: The only thing that I would like to see expanded is the analysis of the impact of the assimilation of the different datasets on the forecast - I only saw one figure 9 which shows the verification of the 21h forecast). It seems that the authors have mainly focused on the verification of the analysis, and the comparison between the run with and without assimilation.

Authors: The way the simulations were performed did not allow the forecast skill of over 21 h forecast to be evaluated, as the simulations were done in an analysis mode (the largest forecast was between the 21 UTC assimilation and 18UTC the next day). To address this comment additional simulations were performed, assimilating data at 18 and 21 UTC and then doing a 48 hour forecast. This was performed only for a period of 10 days and with the MODIS and NNR retrievals only due to computational constraints and because this period had more AOD data. Figure 8 and the following paragraphs were added:

(Section 2.4, 3rd paragraph) Additional simulations were performed for the first 10 days of May to assess the impact of assimilation on forecasts by performing 48 hour unconstrained simulations after each daily 21 UTC analysis.

(Section 3.2, 5th paragraph) An analysis of the impact of assimilation on forecasts starting at 21 UTC is shown in Fig. 8 and 10. When evaluating against PM_{2.5} AQS measurements (Fig. 8), as all forecasts start at the same time, the diurnal cycle modulates the bias and error reductions. For instance, the decreasing trend in fractional error on the 0-4 hour forecast follows the increase in error shown by the non-assimilated model in this portion of the diurnal cycle. PM_{2.5} concentrations show low bias one hour after assimilation, reaching zero values when NASA NNR retrievals are assimilated. Then, the assimilation gradually returns towards concentrations and errors found when no assimilation is performed, in agreement with previous studies (Schwartz et al., 2012). This is also seen in the AQS PM_{2.5} comparison, where assimilation almost never goes back to the non-assimilated model levels (Fig. 3) and fractional error reduction at 18:00 UTC for all stations and days in May is equal to 0.06. After 48 hours there is a slight but positive influence of assimilation for both retrievals (> 0.012 fractional error reduction). These results show that, in the context of operational air quality forecasting, AOD

assimilation with the method developed here can be beneficial for improving the skill of the forecasts for the day after the satellite overpass. As shown earlier, the NNR retrieval assimilation outperforms the MODIS 550nm assimilation for all times for both bias and fractional error.

(Conclusions, 2nd paragraph) 48 hour forecasts starting from an analysis step showed improvements on the aerosol predictions (0.15–0.015 fractional error reductions for the NASA NNR retrieval vs PM2.5), demonstrating the potential of the developed technique for air quality forecasting applications.

Reviewer#2: p.12223 l.14: “successfully” is spelled wrongly.

Authors: Corrected

Reviewer#2: p.12235 ll. 24-28: Where do the dust boundary conditions come from?

Authors: If you mean what data set do they come from, they come from MOZART simulations (already stated in section 2.1). However, MOZART does not compute DUST online, and uses monthly means coming from the CAM model. The following was added to section 2.1: “MOZART uses monthly dust distributions from Community Atmosphere Model (CAM) (Mahowald et al., 2006) calculations, which are also used in this study.”

If you mean from which region do they come from the western and northern boundaries. The following was added in section 3.1: “This overestimation can be traced back to dust aerosol in the chemical boundary conditions coming predominantly from the western and northwestern boundaries.”